HAL open science

High Order C^0-Continuous Galerkin Schemes for High Order PDEs, Conservation of Quadratic Invariants and Application to the Korteweg-de Vries Model

Sebastian Minjeaud, Richard Pasquetti

To cite this version:

Sebastian Minjeaud, Richard Pasquetti. High Order C^0-Continuous Galerkin Schemes for High Order PDEs, Conservation of Quadratic Invariants and Application to the Korteweg-de Vries Model. Journal of Scientific Computing, 2018, 74 (1), pp.491-518. 10.1007/s10915-017-0455-2 . hal-01158007v2

HAL Id: hal-01158007

https://hal.science/hal-01158007v2

Submitted on 5 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

High order C^0-continuous Galerkin schemes for high order PDEs, conservation of quadratic invariants and application to the Korteweg-de Vries model

Sebastian Minjeaud · **Richard Pasquetti**

Received: date / Accepted: date

Abstract We address the Korteweg-de Vries equation as an interesting model of high order partial differential equation, and show that it is possible to develop reliable and effective schemes, in terms of accuracy, computational efficiency, simplicity of implementation and, if required, conservation of the lower invariants, on the basis of a (only) H^1-conformal Galerkin approximation, namely the Spectral Element Method. The proposed approach is *a priori* easily extensible to other partial differential equations and to multidimensional problems.

Keywords Spectral element method; KdV equation; High order derivative approximation; Invariants conservation; IMEX schemes.

Mathematics Subject Classification (2000) 65M60 · 65M70 · 65L06

1 Introduction

As now well known, the spectral element method (SEM) allows a high order approximation of partial differential equations (PDE) and combines the advantages of spectral methods, that is accuracy and rapid convergence, with those of the finite element method (FEM), that is geometrical flexibility, see *e.g.* [13,19,24,31,37]. The SEM is based on a nodal Continuous Galerkin (CG) approach, such that the approximation space contains all C^0 functions whose restriction in each element is a polynomial of degree N. More precisely, in each element, the basis functions are Lagrange polynomials associated to the Gauss-Lobatto-Legendre (GLL) points which are also used as quadrature points to evaluate integrals derived using a weak form of the problem. An

Sebastian Minjeaud
Université Côte d'Azur, CNRS, Inria, LJAD, France.
E-mail: sebastian.minjeaud@unice.fr

Richard Pasquetti
Université Côte d'Azur, CNRS, Inria, LJAD, France.
E-mail: richard.pasquetti@unice.fr

important consequence of this choice for interpolation and integration points is that the resulting mass matrix is diagonal. Thus, the same approach is sometimes called "Gauss-point mass lumped finite element scheme" in the finite element literature [16]. Till now, the fact that the SEM mass matrix is diagonal was mainly exploited to address evolution problems with an explicit time marching, with applications *e.g.* to the wave equation or to the Euler system. Here we use this property to systematically define high order differentiation operators and thus to address dispersive equations.

The efficiency of the SEM is well established for the elliptic or parabolic PDEs, whose solutions are in general smooth. Indeed, many papers describing numerical studies carried out with the SEM are published in the literature, see for instance the papers associated to the ICOSAHOM conference series. Also, following the pioneering code NEKTON developed in the late 80's, parallel softwares like NEKTAR (Imperial College and University of Utah [12]), NEK5000 (Argonne national laboratory [23]) or SPECULOOS (EPFL, Lausanne [20]) make use of the SEM and are routinely used to compute a large variety of problems, *e.g.* , diffusion phenomena, electromagnetism, fluid flows, turbulent flows *etc*... A fine analysis of the efficiency of the SEM, taking into account the computational cost, is *e.g.* carried out in [48], and an interesting parallel between continuous and discontinuous Galerkin method for two-dimensional elliptic problems can be found in [32]. However, when dealing with hyperbolic PDEs, stability problems may be encountered, since, roughly speaking, spectral approximations are centered and much less numerically diffusive than low-order ones. Efficient stabilization techniques are then needed, *e.g.* the variational multiscale method [28], the entropy viscosity method [26] or the spectral vanishing viscosity (SVV) technique [46].

Similarly, extending the SEM approach to dispersive problems has not yet received a great attention and applying the SEM to, *e.g.* , third or higher odd-order equations, remains nontrivial. As a relevant example of such problems we consider the Korteweg-de Vries (KdV) equation, which is well known to point out the dispersive effects that may occur for weakly non-linear water waves or collisionless plasmas, and to provide a simple evidence of the existence of solitons, see *e.g.* [39] and references herein. With $t \in \mathbb{R}^+$ and $x \in \Omega = (x_{\min}, x_{\max})$ for the time and space variables, the KdV PDE may write:

$$\partial_t u + u\partial_x u + \beta \partial_{xxx} u = 0\,, \tag{1}$$

and should be completed with an initial condition $u_0(x) = u(x,0)$, $x \in \Omega$, and *e.g.* periodic boundary conditions. Many numerical methods have been proposed to compute approximate solutions of (1). The existing techniques include finite-difference methods [44,56], spectral methods [45], finite-element methods [4,6,43, 50], and more recently finite volume methods [18,21] or discontinuous Galerkin (DG) methods [7,42,54,55]. In the frame of FEMs, third (or higher) order derivative terms raise some difficulties. When thinking to the standard $\mathbb{P}_1$ FEM, *i.e.* when using a piecewise linear approximation, it is indeed clear that the second order derivative vanishes in each element, so that it is *e.g.* required to use C^1 continuous finite elements or a Petrov-Galerkin approach with C^1 test functions, but at the price of an additional complexity of the algorithm especially when one has in mind high order

approximations or multi-dimensional problems. In this article, we want to show that a high order approximation can be obtained on the basis of C^0 finite elements and propose approaches that remain simple to implement. Moreover, we focus on the preservation of invariants, which is particularly non trivial when high order methods are concerned.

This paper is organized as follows. In Section 2 we first discuss the space discretization for the KdV problem on the basis of the SEM approximation. Two different strategies are introduced to handle the third order derivative term: the first one may be considered as natural to avoid using C^1 finite elements, and the second one as the most general to approximate high order derivative terms with C^0 finite elements. The treatment of the non-linear transport term is also described. The time discretization is discussed in Section 3. Because the preservation of PDE's invariants is generally important, the preservation of the lower invariants of the KdV equation is examined. Here also, two approaches are proposed and carefully detailed. Finally, in Section 4, we experiment the different schemes and algorithms presented in the paper.

2 Space discretization

We first focus on the discretization of the third order derivative term in Section 2.1 and show how to handle boundary conditions in Section 2.2. The discretization of the non-linear transport term is then detailed in Section 2.3.

2.1 Treatment of the third order term

As explained in the introduction, the space discretization is based on the SEM, so that the solution is sought in the space $E_h \subset H^1$ (usual notation is used for the Sobolev space H^1) of all C^0 functions whose restriction in each element is a polynomial of degree N. Then, the scheme is basically obtained using a weak formulation against test functions belonging to the space E_h. Unfortunately, the approximation space E_h is not embedded in the Sobolev space H^2. It raises a difficulty when considering weak formulations associated to a third (or higher) order term, since neither the solution nor the test functions can accept two space derivatives. The basic idea we exploit to overcome this difficulty consists in redefining the first or second order derivative, such that they still belong to the approximation space E_h. This may be interpreted as introducing additional unknowns (belonging to E_h) that allow to decrease the differential order in the weak formulation. Of course, this idea is not new, see *e.g.* [27, 50] for KdV, but with the SEM the key point lies in the fact that the intermediate unknowns can be easily eliminated since the mass matrix is diagonal (recall that the SEM makes use of the same points for interpolation and quadrature rules).

Hereafter, we focus on the space discretization of the third order term, so that we consider the following toy PDE:

$$\partial_t u + \beta \partial_{xxx} u = 0\,, \tag{2}$$

that we associate to an initial condition and to periodic boundary conditions. Periodicity is often assumed for the sake of simplicity: As *e.g.* mentioned in [36], for another variant of the FEM, namely the Local DG (LDG), it is straightforward to enforce other types of boundary conditions as long as the initial-boundary value problem is well posed. To be more precise, in Section 2.2 we however give some details on the resolution of (2) in a finite domain.

Two strategies, that do not suffer from an increase of the number of unknowns and that preserve the high order accuracy of the SEM, are now described.

2.1.1 Strategy A: H^1-approximation of second derivative

The most natural idea consists in introducing $f = \partial_{xx} u$ so that, instead of (2), we address the system

$$\begin{aligned} &\partial_t u + \beta \partial_x f = 0\,, \\ &f = \partial_{xx} u\,. \end{aligned}$$

This system can now be handled by the CG approach and the semi-discrete problem writes: Find $u_h, f_h : \mathbb{R}^+ \to E_h$ such that for any time t

$$\int_{x_{\min}}^{x_{\max}} \partial_t u_h v_h dx + \beta \int_{x_{\min}}^{x_{\max}} \partial_x f_h v_h dx = 0\,, \quad \forall v_h \in E_h\,, \tag{3}$$

$$\int_{x_{\min}}^{x_{\max}} f_h v_h dx + \int_{x_{\min}}^{x_{\max}} \partial_x u_h \partial_x v_h dx = 0\,, \quad \forall v_h \in E_h\,. \tag{4}$$

In algebraic form, one should then consider the system:

$$\begin{aligned} &M\partial_t U + \beta DF = 0, \\ &MF + BU = 0. \end{aligned}$$

The mass matrix M and the matrices D and B write:

$$M_{ij} \approx \int_{x_{\min}}^{x_{\max}} \varphi_i \varphi_j \, dx\,, \quad D_{ij} = \int_{x_{\min}}^{x_{\max}} \varphi_i \partial_x \varphi_j \, dx\,, \quad B_{ij} = \int_{x_{\min}}^{x_{\max}} \partial_x \varphi_i \partial_x \varphi_j \, dx$$

where the set $\{\varphi_i\}$ stand for the usual SEM basis, *i.e.* , in each element the set of Lagrange polynomials associated to the $N+1$ GLL points, and where in each element the GLL quadrature rule is used to compute the integrals. The mass matrix is then diagonal, whereas the integrals defining D and B are computed exactly since the GLL quadrature rule is exact for polynomials of degree $2N-1$. One can eliminate $F = -M^{-1}BU$ to obtain the expression

$$M\partial_t U - \beta DM^{-1}BU = 0,$$

which points out the matrix implementation, in the present CG approach, of the third order derivative term

$$A_1 = -DM^{-1}B.$$

In the general finite element framework, it could be intricate to compute this matrix since it requires to compute the inverse of the mass matrix. Thus, in [49], where the third order derivative is handled in a similar way, it is suggested to lump the mass

matrix, which is not acceptable when high order methods are concerned. Here the breakthrough is that the SEM mass matrix M being diagonal setting up the operator $DM^{-1}B$ is quite easy. Moreover, this property is not specific to the one-dimensional framework.

In the periodic case, an integration by parts shows that $D = -D^t$, *i.e.* D is antisymmetric, so that one also has

$$A_1 = D^t M^{-1} B. \tag{5}$$

In the non periodic case such an expression may be of interest to enforce a boundary value of $\partial_{xx}u$, through an integration by parts of (3), whereas a Neumann condition can be enforced through (4). Of course, boundary terms arise if these conditions are not homogeneous, see Section 2.2.

2.1.2 Strategy B: H^1-approximation of first derivative

Another approach is to define, for any function $u_h \in E_h$, an approximation $u'_h \in E_h$ of its derivative by L^2 projection of $\partial_x u_h$ onto E_h. On the basis of the following problem: Given $u_h \in E_h$, find $u'_h \in E_h$ such that,

$$\int_{x_{\min}}^{x_{\max}} u'_h v_h \, dx = \int_{x_{\min}}^{x_{\max}} \partial_x u_h \, v_h \, dx \,, \quad \forall v_h \in E_h \,,$$

one defines a differentiation operator which can be used to approximate high order derivative terms. Again, this is especially simple when the SEM is concerned, because the SEM matrix is diagonal so that the previous mass matrix problem can be trivially solved. Using the notations introduced previously, one obtains the algebraic expression:

$$MU' = DU \,.$$

Such an approach clearly shows a way to approximate higher order derivative terms: a matrix implementation of derivatives of order p could be $D(M^{-1}D)^{p-1}$. For the equation (2), this approach immediately leads to an expression of a third order derivative algebraic operator

$$A_2 = D(M^{-1}D)^2 \,.$$

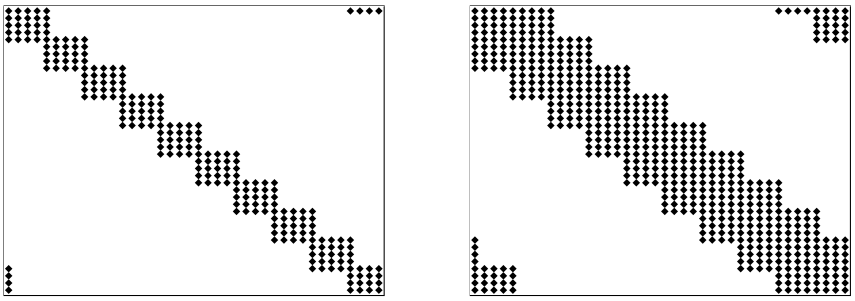
As for strategy A, one may use the antisymmetry of D to obtain equivalent expressions, *e.g.* :

$$A_2 = (D^t M^{-1})^2 D \,, \tag{6}$$

that may be useful in the non periodic case to weakly enforce boundary conditions on $\partial_x u$ or on $\partial_{xx}u$, see Section 2.2.

Coming back to the approximation of $f = \partial_{xx}u$, one has $MF = -D^t M^{-1} DU$, which is directly comparable to the usual weak form, obtained with Strategy A, $MF = -BU$. Especially, both of them allow to enforce Neumann conditions in the non periodic case. The advantage of Strategy A is that matrix B can be set up by assembling elementary matrices, contrarily to the algebraic operator that we have introduced. Hence, matrix B is mainly composed of overlapping square blocks of

size $(N+1)$, the overlaps being scalars, whereas one easily checks that for matrix $D^tM^{-1}D$ they are of size $(2N+1)$, the overlaps being now square matrices of size $(N+1)$, see Fig. 1. Then, matrix A_1 and matrix A_2 are mainly composed of overlapping square blocks of size $(2N+1)$ and $(3N+1)$, respectively.



Matrix B Matrix $D^tM^{-1}D$

Fig. 1 Stencil of matrices. $N = 4$, 10 elements.

It should be mentioned that the kernel of the matrix D, and consequently the kernel of the matrix $D^tM^{-1}D$, is of dimension 2. Because of the periodic boundary conditions, as the kernel of the matrix B it contains the constant vector, which is associated in each element to the Legendre polynomial $L_0(x)(=1)$, but also the vector that in each element is associated to the Legendre polynomial $L_N(x)$. This directly results from the definition of the GLL points, where $L_N(x)$ finds his extrema. More precisely, in each element and up to scaling factors: If N is even the eigenfunction coincides with $L_N(x)$, whereas if N is odd it alternatively coincides with $\pm L_N(x)$, so that in both case the eigenfunction is in E_h. Since we consider a time evolution problem, this does not impact the stability of the algorithm. Finally, we have numerically checked that the condition number of the matrices B and $D^tM^{-1}D$, defined as the ratio of the largest eigenvalue over the smallest non zero one, have similar behaviours: K^2N^3, K being the number of elements, see Fig. 2. Consequently, the computational costs are comparable and one may think that a preconditionner efficient for B will be efficient for $D^tM^{-1}D$, but in the frame of a one-dimensional equation like KdV this is of limited interest.

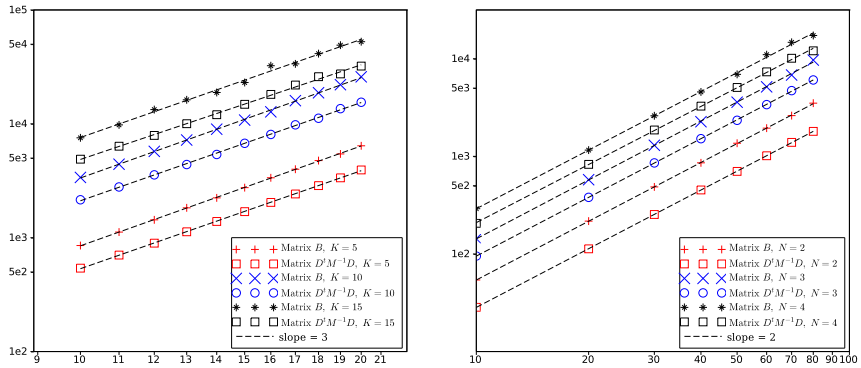
Strategy B is equivalent to consider the problem: Find u_h, f_h and $g_h : \mathbb{R}^+ \to E_h$, such that:

$$\int_{x_{\min}}^{x_{\max}} \partial_t u_h v_h dx + \beta \int_{x_{\min}}^{x_{\max}} \partial_x f_h v_h dx = 0\,, \quad \forall v_h \in E_h\,, \tag{7}$$

$$\int_{x_{\min}}^{x_{\max}} f_h v_h dx = -\int_{x_{\min}}^{x_{\max}} g_h \partial_x v_h dx\,, \quad \forall v_h \in E_h\,, \tag{8}$$

$$\int_{x_{\min}}^{x_{\max}} g_h v_h dx = \int_{x_{\min}}^{x_{\max}} \partial_x u_h v_h dx\,, \quad \forall v_h \in E_h\,. \tag{9}$$

Such an approach is in fact typical of the DG method, see *e.g.* the formulation considered in [52,53] for the LDG approximation of (2). Recall that LDG is called "local",



As a function of N (polynomial degree) As a function of K (number of elements)

Fig. 2 Condition number of matrices B and $D^tM^{-1}D$

because the elimination of the additional variables is made locally, *i.e.* in each element. Consequently, just like in our SEM approach, no global mass matrix has to be inverted. Of course, LDG also requires to set up *adhoc* numerical fluxes to obtain a well posed discrete problem. Let us also mention the recent work [22], where the LDG technique is implemented both in space and in time, in order to avoid a severe constraint on the time step. As developed in Section 3.1, to avoid such a constraint we treat implicitly the third order derivative term.

Here one can state that:

Lemma 1 *If using strategy B, the semi-discrete SEM approximation of* (2) *yields the exact conservation of the discrete energy.*

Proof With $v_h = u_h$ in (7), $v_h = g_h$ in (8) and $v_h = f_h$ in (9), one obtains:

$$\frac{1}{2}\int_{x_{\min}}^{x_{\max}} \partial_t u_h^2\, dx = \beta \int_{x_{\min}}^{x_{\max}} g_h \partial_x g_h\, dx = \frac{\beta}{2}[g_h^2]_{x_{\min}}^{x_{\max}} = 0$$

by periodicity. These equalities still hold when using the GLL quadrature rule: Indeed, in each element the derivative of a polynomial of degree N is exactly computed and the GLL quadrature is exact for polynomials of degree $2N-1$.

The result can also be derived from the algebraic system. Indeed, since the matrix A_2 is antisymmetric, when using Strategy B we have:

$$\langle A_2 U, U\rangle = 0 \Rightarrow \partial_t \langle MU, U\rangle = 0\,.$$

2.2 The non periodic case

For the sake of simplicity the periodic case has till now been considered. We show in this Section that the non periodic one can also be easily addressed. To be more

precise, we assume that a Dirichlet condition is imposed at $x_{\min}$ and that $\partial_x u$ and $\partial_{xx} u$ are given at $x_{\max}$, *i.e.*

$$u(x_{\min}) = b_-, \quad \partial_x u(x_{\max}) = b'_+, \quad \text{and} \quad \partial_{xx} u(x_{\max}) = b''_+,$$

where b_-, b'_+ and b''_+ are given data, that may depend on the time variable. In the SEM (and FEM) framework, Dirichlet conditions are usually addressed strongly. This means that the space E_h is now restricted to the functions which additionally vanish at $x_{\min}$. It raises a difficulty for the discrete variable f_h (and g_h in strategy B) for which no Dirichlet boundary condition is imposed at $x_{\min}$. A unique grid solution consists in keeping the GLL points approximation while constraining the value of $f_h(x_{\min})$ using a polynomial extrapolation from the other values of f_h in the element that contains $x_{\min}$. This is natural here since it means that in the left hand side boundary element f_h is a polynomial of degree $N-1$. Indeed, in each element $\partial_{xx} u_h \in \mathbb{P}_{N-2}$, and looking for $f_h \in \mathbb{P}_N$ relies on the need of the C_0 continuity or on the need to impose a boundary condition on $\partial_{xx} u$, as on the right hand side element. In the left hand side element, there is clearly only one constraint to impose (the C_0 continuity at its right interface). Note that this approach using extrapolation can be easily extended to multidimensional problems.

2.2.1 Strategy A

To enforce the considered boundary conditions the discrete system is rewritten:

$$\int_{x_{\min}}^{x_{\max}} \partial_t u_h v_h dx - \beta \int_{x_{\min}}^{x_{\max}} f_h \partial_x v_h dx = -\beta b''_+ v_h(x_{\max}), \quad \forall v_h \in E_h,$$
$$\int_{x_{\min}}^{x_{\max}} f_h v_h dx + \int_{x_{\min}}^{x_{\max}} \partial_x u_h \partial_x v_h dx = b'_+ v_h(x_{\max}), \quad \forall v_h \in E_h.$$

To set up the algebraic system, let us split the unknown vector U in $U = U_0 + U_1$, where U_1 takes into account the Dirichlet boundary value $u(x_{\min}) = b_-$, *i.e.* all the components of U_1 are zero except one which equals b_-. Similarly, let us state that $F = F_0 + F_1$, where the non null value of F_1 is computed by polynomial extrapolation. One may write $F = EF_0$. The matrix E is an easily identifiable extrapolation matrix: Only one line differs from the identity matrix, and its entries are computed, for the reference element $(-1,1)$, using the usual expressions of the Lagrange polynomials. Additionally, we introduce the diagonal matrix R such that $U_0 = RU$ (and $F_0 = RF$), *i.e.* R is a diagonal matrix with diagonal terms equal to 1 except one term set to 0. Then one can set up the discrete system

$$M\partial_t U_0 - \beta R D^t E F_0 = S_1,$$
$$MF_0 + RBU_0 = S_2 - RBU_1.$$

where S_1 is associated to the conditions $\partial_{xx} u_h = b''_+$ and S_2 to the condition $\partial_x u_h = b'_+$. Note that for the sake of simplicity the notations of the mass and differentiation matrices have not been changed. In our notations, their dimensions equal the total number of grid points. Note also that two equations of the system resume to $0 = 0$, to

be replaced by (i) the Dirichlet value $u_h(x_{\min}) = b_-$ and (ii) the extrapolation equation that gives $f_h(x_{\min})$. By elimination of the vector F_0 one obtains:

$$M\partial_t U_0 + \beta RD^t EM^{-1} RBU_0 = S\,, \quad S = S_1 + \beta RD^t EM^{-1}(S_2 - RBU_1)\,.$$

Clearly, this system is no longer homogeneous but can be handled similarly to the one obtained when assuming periodicity, see (5).

2.2.2 Strategy B

To include the considered boundary conditions, the discrete system is rewritten:

$$\begin{aligned}
&\int_{x_{\min}}^{x_{\max}} \partial_t u_h v_h dx - \beta \int_{x_{\min}}^{x_{\max}} f_h \partial_x v_h dx = -\beta b''_+ v_h(x_{x_{\max}})\,, \quad \forall v_h \in E_h\,,\\
&\int_{x_{\min}}^{x_{\max}} f_h v_h dx = -\int_{x_{\min}}^{x_{\max}} g_h \partial_x v_h dx + b'_+ v_h(x_{x_{\max}})\,, \quad \forall v_h \in E_h\,,\\
&\int_{x_{\min}}^{x_{\max}} g_h v_h dx = \int_{x_{\min}}^{x_{\max}} \partial_x u_h v_h dx\,, \quad \forall v_h \in E_h\,.
\end{aligned}$$

By proceeding as for strategy A one obtains the algebraic system:

$$\begin{aligned}
&M\partial_t U_0 - \beta RD^t EF_0 = S_1,\\
&MF_0 + RD^t EG_0 = S_2,\\
&MG_0 - RDU_0 = RDU_1\,.
\end{aligned}$$

By elimination of F_0 and G_0 one obtains:

$$M\partial_t U_0 + \beta (RD^t EM^{-1})^2 RDU_0 = S\,, \quad S = S_1 + \beta RD^t EM^{-1} S_2 - \beta (RD^t EM^{-1})^2 RDU_1\,.$$

As for strategy A, the obtained system is non homogeneous but similar to the one obtained for the periodic problem, see (6).

2.3 Treatment of the transport term

The transport term is handled by adding to the discrete formulation the following term:

$$\int_{x_{\min}}^{x_{\max}} u_h \partial_x u_h v_h \, dx, \tag{10}$$

where v_h is the test function. Let us remark that, on each element, $u_h \partial_x u_h v_h$ is a polynomial function of degree $3N-1$. Thus, when using the GLL quadrature associated to the approximation degree N (which is exact only for polynomial functions up to degree $2N-1$), the term (10) is not exactly computed. In numerical simulations, we observe that it can lead to spurious oscillations, so that the use of a more accurate quadrature rule is desirable to compute the term (10). Also, an exact integration has an interesting theoretical consequence: in this case Lemma 1 still holds since the formula

$$\int_{x_{\min}}^{x_{\max}} u_h^2 \partial_x u_h \, dx = \frac{1}{3}[u_h^3]_{x_{\min}}^{x_{\max}} = 0,$$

by periodicity, is exactly computed. Thus, if using Strategy B, the L^2 norm of the semi-discrete solution is exactly conserved during the time evolution.

In the frame of the SEM, quadrature rules as accurate as required can be obtained by using a GLL grid associated to a polynomial degree $M > N$. This finer GLL grid is only introduced on the reference element and an extension operator from the initial GLL grid (associated to the degree N) to the finer one (associated to the degree M) is set up on the reference element. The quadrature rule associated to the finer GLL grid is of order $2M-1$, so that an exact integration of the term (10) can be obtained with M such that $2M-1 \geqslant 3N-1$, *i.e.* $M \geqslant 3N/2$. Nearly all the results presented in Section 4 are obtained with this "overintegration" strategy. However, as outlined in [33,40], using $M = N+1$ or $M = N+2$ may be sufficient to avoid the spurious oscillations.

To end this section, we emphasize the fact that a viscous stabilization technique is not required unless very coarse grids are used. This is due to the fact that the third order derivative term provides a regularization, see *e.g.* [30], so that the solution of KdV is smooth, provided that the initial data is smooth. On very coarse mesh, the SVV technique, see Appendix A, may be of interest. Indeed, in the frame of spectral methods this technique is known to be efficient for the inviscid Burgers equation, which formally can be obtained from the KdV equation in the limit $\beta = 0$. Even if this discussion is beyond the scope of the present paper, it is however important to mention that the use of a viscous stabilization may be criticized since it is known [35] (see also [27]) that diffusive and dispersive regularizations of the inviscid Burgers equation lead to different solutions when the regularization parameter tends to 0. For comparisons, in Section 4.6, we carry out computations using the SVV technique and point out the influence of such a stabilization.

3 Time discretization and conservation of invariants

For the sake of computational efficiency, we propose to handle the non-linear term $\mathscr{N}(u) = u\partial_x u$ explicitly, whereas for the sake of numerical stability, we propose an implicit treatment of the linear term $\mathscr{L}(u) = \beta\partial_{xxx}u$. This indeed allows a time-independent algebra, *i.e.* the algebraic operator associated to the linear term is only inverted once in a preprocessing stage. Moreover, with this choice, one can expect the usual Courant-Friedrichs-Lewy (CFL) stability condition. To obtain a high order approximation in coherence with the SEM, we propose to use an implicit-explicit (IMEX) Runge Kutta (RK) scheme. This is explained in Section 3.1. However, although IMEX schemes ensure the preservation of all linear invariants (as the first invariant associated to the KdV equation), they do not allow to preserve all quadratic invariants (as the second one, see (11)). We remedy this problem, in Section 3.2, by proposing an additional step which ensures the conservation of a specific quadratic invariant while preserving the accuracy order of the time scheme.

3.1 Implicit-explicit Runge Kutta scheme

The IMEX schemes were first developed in the 90's [2] and details may now be found in several papers, see *e.g.* [41]. They combine an Explicit RK (ERK) scheme for the operator $\mathscr{N}$ and an Implicit RK (IRK) scheme for the operator $\mathscr{L}$. Denoting by τ the time step, a s-stages IMEX scheme for the ODE $\partial_t u = \mathscr{N}(u) + \mathscr{L}(u)$, may read as follows

$$\begin{cases} u_{n,i} = u_n + \tau \sum_{j=1}^{s} \left(\hat{a}_{ij} \mathscr{N}_j + a_{ij} \mathscr{L}_j \right), \quad \forall i = 1,..,s, \\ u_{n+1} = u_n + \tau \sum_{j=1}^{s} \left(\hat{b}_j \mathscr{N}_j + b_j \mathscr{L}_j \right). \end{cases}$$

where the $\mathscr{N}_j \equiv \mathscr{N}(u_{n,j})$ and $\mathscr{L}_j \equiv \mathscr{L}(u_{n,j})$ are the values of $\mathscr{N}(u)$ and $\mathscr{L}(u)$ at the intermediate RK steps. The two schemes are defined by the coefficients $\hat{A} = (\hat{a}_{ij})$, $\hat{b} = (\hat{b}_j)$ and $A = (a_{ij})$, $b = (b_j)$, $1 \leqslant i, j \leqslant s$, which are classically gathered in a Butcher tableau, see *e.g.* [9]. The coefficients in $\hat{A}$ should vanish when $j \geqslant i$ since they define an explicit scheme. Concerning the implicit part, IMEX schemes generally make use of Diagonally IRK (DIRK) schemes, that is $a_{ij} = 0$ if $j > i$. Consequently, the intermediate unknowns $u_{n,i}$ can be successively computed by "inverting" the operators $Id - \tau a_{ii} \mathscr{L}$, for $i = 1,..,s$ (Id, for identity operator). Moreover, it is usual to assume that the normalized intermediate times $c_i = \sum_{j=1}^{s} a_{ij} = \sum_{j=1}^{s} \hat{a}_{ij}$ of the two schemes coincide and also the coefficients in $\hat{b}$ and b are often the same (in that case, the order of the resulting IMEX scheme is exactly the minimum of the order of its two constituting schemes).

An IMEX scheme preserves all linear invariants but, as an ERK scheme, it cannot preserve all quadratic invariants. However, it is possible to focus on a specific quadratic invariant and ensure its conservation (see Section 3.2).

In this article, we mainly use the IMEX scheme ARS (2,3,3) [3], characterized by 2 implicit steps, 3 explicit ones and which is globally third order accurate. Using Butcher's like notation, the coefficients of ARS(2,3,3) are provided in Table 1, where $\gamma = (3+\sqrt{3})/6$.

$$\begin{array}{c|c||c} c & \hat{A} & A \\ \hline & \hat{b} & b \end{array} \quad \longleftrightarrow \quad \begin{array}{c|ccc||ccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \gamma & \gamma & 0 & 0 & 0 & \gamma & 0 \\ 1-\gamma & \gamma-1 & 2(1-\gamma) & 0 & 0 & 1-2\gamma & \gamma \\ \hline & 0 & 1/2 & 1/2 & 0 & 1/2 & 1/2 \end{array}$$

Table 1 IMEX coefficients for ARS(2,3,3). The ERK scheme is defined at left of the vertical double bar and the DIRK scheme at right . The normalized intermediate times c_i are given in left column.

3.2 Preservation of invariants

As *e.g.* explained in [25], the KdV equation is characterized by an infinity of invariants. The three lowest ones write:

$$C_1 = \int_{x_{\min}}^{x_{\max}} u\,dx\,, \quad C_2 = \int_{x_{\min}}^{x_{\max}} u^2\,dx\,, \quad C_3 = \int_{x_{\min}}^{x_{\max}} \left(u^3 - 3\beta(\partial_x u)^2\right)\,dx\,. \quad (11)$$

As a direct consequence of the weak formulation, one has:

Proposition 1 *Using any consistent approximation of the operator ∂_t, the spectral element schemes described in Section 2.1 preserve the mass invariant C_1.*

Proof Using a constant for test-function one obtains:

$$\int_{x_{\min}}^{x_{\max}} \partial_t u_h\,dx = -\int_{x_{\min}}^{x_{\max}} (u_h \partial_x u_h + \beta\,\partial_x f_h)\,dx\,.$$

Since in each element the SEM approximation of ∂_x is exact for any polynomial of degree N and since the GLL quadrature is exact for polynomials of degree $2N-1$, the discrete SEM computation exactly yields

$$\frac{d}{dt}\left(\int_{x_{\min}}^{x_{\max}} u_h\,dx\right) = -\left[\frac{u_h^2}{2} + \beta f_h\right]_{x_{\min}}^{x_{\max}} = 0\,.$$

Note that this holds with or without overintegration of the non-linear term.

Such a result is rather satisfactory, since many of the schemes described in the literature on the KdV equation do not show the same property.

The time discretization plays a key role in the preservation of the energy invariant C_2. To ensure its conservation, a first possibility is to make use of a scheme especially designed to this end. An example of such a scheme is the standard Crank-Nicolson (CN) scheme. This approach is retained *e.g.* in the recent paper [55] where the CN scheme is used in conjunction with a specific DG approximation designed by tuning the DG parameters to preserve the two first invariants.

One can state the following:

Proposition 2 *If (i) the integral associated to the non-linear term is exactly computed and if (ii) the third order term is formulated using Strategy B, then the CN scheme preserves the two first invariants.*

Proof As soon as (i) is verified, this is a direct consequence of Lemma 1 and of the fact that for CN:

$$\partial_t(u^2) = 2u\partial_t u \approx (u_{n+1} + u_n)\frac{u_{n+1} - u_n}{\tau} = \frac{1}{\tau}(u_{n+1}^2 - u_n^2)$$

where u_n, u_{n+1} are the numerical solutions at time t_n, t_{n+1}, $\tau = t_{n+1} - t_n$ being the time step.

Using this approach shows however two drawbacks: (i) loss of time accuracy, since the CN scheme is only second order accurate, and (ii) increase of the computational cost, since a non-linear solve is needed at each time-step. Going back to RK methods, it is known since [17] that a RK method preserves all quadratic invariants if and only if its coefficients satisfy $b_i a_{ij} + b_j a_{ji} = b_i b_j$ for all $1 \leqslant i, j \leqslant s$, see also *e.g.* [15,29]. These conditions impose strong requirements on the coefficients and are only satisfied by some IRK methods. Nevertheless, even if ERK or IMEX RK cannot preserve all quadratic invariants, it is possible to focus on the preservation of a specific invariants and ensure it by projection at the end of each time step while preserving the accuracy of the overall scheme. Hereafter, two possibilities are discussed.

- A first possibility is to complete the IMEX RK scheme by a L^2 projection of the RK solution onto the hypersurface associated to the constraints. This can be achieved by introducing time dependent Lagrange multipliers. As shown in Section 4, the two first invariants can then be exactly preserved. Because one can derive an exact expression of the Lagrange multipliers the computational efficiency is maintained. Details on the implementation of this approach, and on the fact that the time accuracy is preserved, are provided in Section 3.2.1.
- A second possibility is to proceed by interpolation / extrapolation, at each time-step, of the solutions obtained by using two different RK schemes. One can also look at that as a non-orthogonal projection [8,10,11]. Then, since both schemes preserve the first invariant, it is only required to compute an interpolation / extrapolation time dependent factor that allows to preserve the second invariant. If using such an approach, the computational cost is *a priori* twice greater, but this is no longer true with an embedded RK IMEX scheme. In this case, one only modifies the b_i so that the $u_{n,i}$ are the same (see Section 3.1). Moreover, even if the embedded scheme is of order $q < p$, the accuracy may be preserved. For instance, from ARS(2,3,3) one obtains a lower order scheme simply by modifying the last line of Table 1. Because for consistency reason the sum of these coefficients should equal 1, in Section 4 we simply use the pair $\{1/4, 3/4\}$ instead of $\{1/2, 1/2\}$, for both the ERK and the DIRK parts, which yields a first order IMEX scheme. Details are given in Section 3.2.2.

The two algorithms described hereafter are valid for any RK schemes, are not specific to the KdV equation and are not restricted to one dimensional problems. They however focus on the preservation of a specific quadratic invariant. Up to our knowledge, they have never been used for KdV. They are numerically compared in Sections 4.1 and 4.2.

3.2.1 Algorithm 1: Invariants preservation by projection

At each time step, we proceed by discrete L^2 projection of the IMEX solution onto the manifold associated to the discrete versions of the two first invariants. Starting from the discrete solution $u_h(x,t_n)$ at time t_n, we first compute the grid-point values y_i using the RK scheme and then define $u_h(x,t_{n+1})$ by orthogonal projection.

Denoting w_i the quadrature coefficient associated to the grid-point x_i, $1 \leqslant i \leqslant d$, we define $C_1 = \sum_i w_i u_h(x_i,t_n)$ and $C_2 = \sum_i w_i (u_h(x_i,t_n))^2$. The optimization problem

then writes: Find the δy_i that minimize the functional $F(\delta y_1, \cdots, \delta y_d) = \sum_i w_i \delta y_i^2$, such that $\sum_i w_i(y_i + \delta y_i) = C_1$ and $\sum_i w_i(y_i + \delta y_i)^2 = C_2$. The discrete solution at time t_{n+1} is then set as $u_h(x_i, t_{n+1}) \equiv u_i = y_i + \delta y_i$.

To transform the constrained optimization problem into a non constrained one, we introduce the functional:

$$L(\delta y_1, \cdots, \delta y_d, \lambda_1, \lambda_2) = \\ \sum_i w_i \delta y_i^2 + \lambda_1 \Big(\sum_i w_i(y_i + \delta y_i) - C_1 \Big) + \lambda_2 \Big(\sum_i w_i(y_i + \delta y_i)^2 - C_2 \Big)$$

and the problem is now to find the δy_i together with the Lagrange multipliers λ_1 and λ_2. At the optimum of this saddle point problem, the gradient of the functional vanishes. This yields:

$$\partial_{\delta y_i} L = 2 w_i \delta y_i + \lambda_1 w_i + 2\lambda_2 w_i (y_i + \delta y_i) = 0 \tag{12}$$

$$\partial_{\lambda_1} L = \sum_i w_i (y_i + \delta y_i) - C_1 = 0 \tag{13}$$

$$\partial_{\lambda_2} L = \sum_i w_i (y_i + \delta y_i)^2 - C_2 = 0 \tag{14}$$

so that, from (12)

$$\delta y_i = -\frac{\lambda_2}{1+\lambda_2} y_i - \frac{\lambda_1}{2(1+\lambda_2)} \quad \text{and} \quad u_i = \frac{1}{1+\lambda_2} y_i - \frac{\lambda_1}{2(1+\lambda_2)} . \tag{15}$$

It remains to plug the last expression in (13) and (14). Using the notation $S_y = \sum_i w_i y_i$, $S_{y^2} = \sum_i w_i y_i^2$ and $S_1 = \sum_i w_i = x_{\max} - x_{\min}$, one obtains:

$$S_y - S_1 \frac{\lambda_1}{2} - C_1(1+\lambda_2) = 0 \tag{16}$$

$$S_{y^2} - S_y \lambda_1 + S_1 \frac{\lambda_1^2}{4} - C_2(1+\lambda_2)^2 = 0$$

This non-linear system in λ_1 and λ_2 can be easily solved. With $(1+\lambda_2)$ from (16), one obtains that λ_1 solves:

$$\frac{1}{4}(S_1^2 - \alpha S_1)\lambda_1^2 - (S_1 S_y + \alpha S_y)\lambda_1 + S_y^2 - \alpha S_{y^2} = 0 \tag{17}$$

where $\alpha = C_1^2 / C_2 > 0$. One can compute λ_1 from this equation and then obtain λ_2 from (16). Note that since $\alpha \approx S_y^2 / S_{y^2}$, the negative discriminant is relevant, because in case of equality this yields $\lambda_1 = 0$. Once knowing the Lagrange multipliers, from (15) one computes the u_i.

Now we show the following:

Proposition 3 *Algorithm 1 preserves the approximation order of the RK scheme.*

Proof Assuming we use a RK scheme of order p ($p = 3$ in our examples) to compute the y_i, $1 \leqslant i \leqslant d$, then the local truncation error, *i.e.* the gap between the y_i and the exact solution obtained at time t_{n+1} starting from $u_h(x,t_n)$, is $O(\tau^{p+1})$. We can readily deduce that $C_1 - S_y$ and $C_2 - S_{y^2}$ are also $O(\tau^{p+1})$. Next we obtain that

$$S_1^2 - \alpha S_1 = O(1), \quad S_1 S_y + \alpha S_y = O(1) \quad \text{and} \quad S_y^2 - \alpha S_{y^2} = O(\tau^{p+1}).$$

Equation (17) then shows that λ_1 is $O(\tau^{p+1})$ and we deduce from equation (16) that λ_2 is also $O(\tau^{p+1})$. Taking into account the fact that $|\lambda_1|, |\lambda_2| \ll 1$, from (15) one has:

$$\delta y_i \approx -\lambda_2 y_i - \frac{\lambda_1}{2}$$

so that the correction behaves like the Lagrange multipliers. As a result, the correction by the δy_i is consistent with the accuracy of the RK scheme.

It should be noticed that more than two invariants can be considered, but certainly at the price of a non-explicit formulation of the Lagrange multipliers.

3.2.2 Algorithm 2: Invariants preservation by interpolation / extrapolation

Here we assume to have at hand two RK IMEX schemes. Knowing the numerical solution at time t_n, $u_h(x,t_n)$, we compute the grid-point values y_i and z_i, $1 \leqslant i \leqslant d$, at time t_{n+1} with the two RK IMEX schemes. Because each scheme preserves the invariant C_1, any linear combination of the two solutions will be C_1-invariant. The goal is then to define an interpolation / extrapolation factor λ, such that C_2 is preserved. Thus, at time t_{n+1}, we define

$$u_i = (1-\lambda) y_i + \lambda z_i\,, \quad \sum_i w_i u_i^2 = C_2 \tag{18}$$

where again $C_2 = \sum_i w_i (u_h(x_i,t_n))^2$, and w_i are the quadrature coefficients associated to the grid-points x_i.

Using notations similar to those previously introduced, one easily checks that λ should solve:

$$S_{(z-y)^2} \lambda^2 + 2 S_{y(z-y)} \lambda + S_{y^2} - C_2 = 0\,. \tag{19}$$

The (reduced) discriminant writes, $\Delta = S_{y(z-y)}^2 + S_{(z-y)^2}(C_2 - S_{y^2})$. It is positive as soon as $C_2 \geqslant S_{y^2}$, *i.e.* if the y-scheme is dissipative. In case of equality, the coefficient λ should vanishes, which means that the relevant solution is obtained by using the sign of the quantity $S_{y(z-y)}$ against the discriminant.

As for Algorithm 1, one has the following less intuitive result:

Proposition 4 *Algorithm 2 generally preserves the approximation order of the RK scheme.*

Proof Let us assume that the y-scheme is of order p and the z-scheme of order $q < p$. By definition, for any grid-point we can write

$$y_i = \bar{u}_i + O(\tau^{p+1}), \quad z_i = \bar{u}_i + \psi_i \tau^{q+1} + O(\tau^{q+2}),$$

where $\bar{u}_i$ is the exact solution, obtained at point x_i and time t_{n+1}, when starting from $u_h(x,t_n)$. Hence, coming back to the coefficients of equation (19) we have

$$S_{(z-y)^2} = O(\tau^{2q+2}), \quad S_{y(z-y)} = S_{\bar{u}\psi}\tau^{q+1} + O(\tau^{q+2}), \quad S_{y^2} - C_2 = O(\tau^{p+1}).$$

If we assume that $S_{\bar{u}\psi} \neq 0$, then the discriminant of (19) is positive, at least for τ small enough, and we obtain that λ is $O(\tau^{p-q})$. Hence, from the expression of u_i in (18), the accuracy of the y-scheme is preserved.

A similar result was previously provided for ERK schemes in [10, Theorem 3.1].

4 Numerical experiments

The numerical examples provided in this section are considered as standard benchmarks for KdV solvers. For the two first examples, in Sections 4.1 and 4.2, we refer to [25] or *e.g.* to [1,14,47,57] when using the FEM. The third example, in Section 4.3, goes back to the pioneering paper [56], see also *e.g.* [58,18]. In Section 4.4, as in [18] we investigate the long-time behaviour of the scheme considering a one-soliton solution. In Section 4.5 we revisit a non periodic problem, *i.e.* involving boundary conditions, investigated in [54].

Unless otherwise specified, the third order derivative is approximated on the basis of Strategy A, see Section 2.1.1, and an exact integration of the non-linear term is achieved by using the GLL grid associated to the polynomial approximation degree $M \geqslant 3N/2$, see Section 2.3. In the sequel, we use values of N up to 5, so that we take $M = 8$. The time discretization is based on a globally third order accurate IMEX scheme: ARS (2,3,3) [3] (see Section 3.1).

We point out some differences between the two strategies presented in this paper and also discuss the interest of the overintegration and of the SVV technique in Section 4.6.

4.1 Test-case 1

This example describes the interaction between two solitons. Because an exact solution is known for the open domain $\Omega \equiv \mathbb{R}$, it allows to carry out an accuracy study.

As described *e.g.* in [25], for $\beta = 4.84\,10^{-4}$ the following analytic expression solves the KdV equation:

$$u_{ex}(x,t) = 12\beta(LogF)_{xx},$$

with $F = 1 + e^{\eta_1} + e^{\eta_2} + \alpha e^{(\eta_1+\eta_2)}$ and $\eta_i = \alpha_i x - \alpha_i^3 \beta t + b_i$, $i = 1,2$, the values of the coefficients α_1, α_2, b_1 and b_2 being defined by the relationships

$$\alpha = \left(\frac{\alpha_1 - \alpha_2}{\alpha_1 + \alpha_2}\right)^2, \ \alpha_1 = \sqrt{0.3/\alpha}, \ \alpha_2 = \sqrt{0.1/\alpha}, \ b_1 = -0.48\alpha_1, \ b_2 = -1.07\alpha_2.$$

Computations have been carried out for $x \in (-1,4)$ and $t \in (0,6.3)$, assuming periodicity and using for initial condition the exact value at $t = 0$. In the considered time interval, on an animation one clearly observes the propagation of the two solitons of magnitude 0.9 and 0.3, their crossing and then their reformation. Fig. 3 shows this time evolution in the (x,t)-plane.

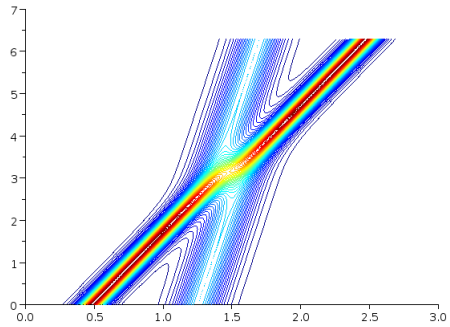


Fig. 3 Test-case 1: Contour levels of the numerical solution in the (x,t)-plane. Computation done with $K = 300$ elements, a polynomial degree $N = 5$ and a time step $\tau = 9\ 10^{-5}$.

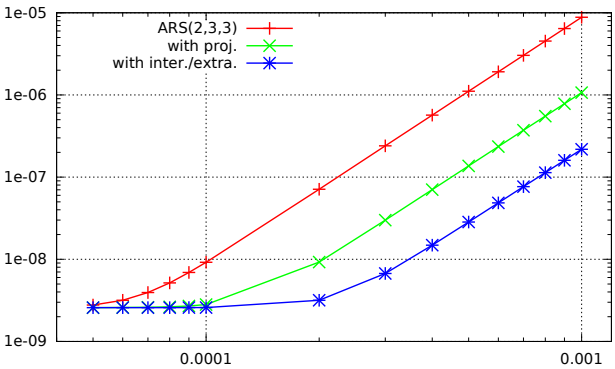


Fig. 4 Test-case 1: Time accuracy for the ARS(2,3,3) scheme, in its basic version and when associated to an additional correcting step. L^2-error as a function of the time step.

We first check if the third order time accuracy of the RK IMEX scheme is obtained. To this end, computations have been carried out with the fine grid obtained for $K = 300$ and $N = 5$, using different time-steps. Moreover, in Section 3.2 it is proved that the accuracy of the IMEX scheme is preserved when completing it with an additional step, in view of preserving the two first invariants. To check that, computations have also been carried out using (i) the projection and (ii) the interpola-

tion/extrapolation techniques. Fig. 4 shows the L^2 norm of the gap between the numerical and exact solutions. Before the saturation associated to the error in space, one clearly observes that ARS(2,3,3) is third order accurate, even if associated with an additional correcting step. In the present example, it turns out that the accuracy is improved when using a correcting step, and that the best accuracy is obtained with the interpolation/extrapolation technique. Hereafter, this is the approach used to check the space accuracy.

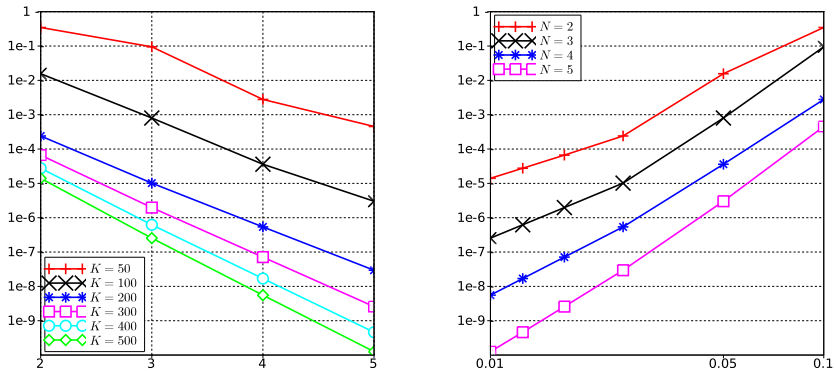


Fig. 5 Test-case 1: L^2-error as a function of the polynomial degree N (at left) and of the element size h (at right)

To carry out the space accuracy study, we have considered polynomial approximation degrees such that $2 \leqslant N \leqslant 5$ and the following numbers of elements, $K \in \{50, 100, 200, 300, 400, 500\}$. To make negligible the time-stepping error, we have used the small time-step value $\tau = 5\ 10^{-5}$. Results are provided in Fig. 5. As expected, one clearly observes, in Fig. 5(left) the exponentially fast decrease of the error with respect to the polynomial degree N and in Fig. 5(right) an algebraic convergence with respect to the element size h of order about $N+1$, see Table 2, which is optimal. It should however be mentioned that:

- Such accuracy results cannot be obtained for $\Omega = (0,4)$, as considered in [1, 25,57]. In these papers, the listed errors remain indeed greater than the value of $u_{ex}(0,0)$, *i.e.* , around 10^{-6}. As a result, with this computational domain one observes a saturation in the decrease of the error. The phenomenon is overcome when using $\Omega = (-1,4)$, because at both end sides, u_{ex}, which is defined for $\Omega \equiv \mathbb{R}$, is then smaller than the round off error.

- To avoid an artifact due to a superconvergence associated to the quadrature rule, the L^2 errors have been computed by using the overintegration technique, *i.e.* in each elements the GLL points obtained for $N = 8$ have been used as quadrature points. Some convergence rates of Fig. 5 may even seem greater than $N+1$. This is due to the fact that for too rough discretizations, *e.g.* , $N = 2$ and $K \leqslant 200$, the asymptotic regime

is not yet reached. We are here in the underresolved case considered in Section 4.6. If using only the results obtained for $h \leqslant 0.025$, *i.e.* $K \geqslant 200$, linear regressions (carried out for the logarithms of the L^2 error and of the grid size) provide the convergence rates given in Table 2.

N	2	3	4	5
Conv. rate	3.099	4.026	4.996	5.962

Table 2 Test-case 1: L^2 error convergence rates for $N \in \{2,3,4,5\}$.

4.2 Test-case 2

Starting from a Gaussian as initial condition, interesting solutions may be obtained. They crucially depend on the value of the β parameter with respect to a critical value β_c. For $\beta \ll \beta_c$, the initial condition splits into a series of stiff solitons, which number depends on the value of β. We use this example to show that even for very stiff solutions, the high order approximation may yield satisfactory results. Moreover, as *e.g.* in [25], we check on this example the conservation of the 3 lowest invariants, see equation (11).

As in [25,57] we solve the KdV equation for $t \in (0,12.5)$ and in the computational domain $\Omega = (-15,15)$. For initial condition we use $u(x,0) = \exp(-x^2)$, so that the critical β is $\beta_c = 0.0625$ [5]. We choose the value $\beta = 10^{-3} \ll \beta_c$, so that one expects the formation of 9 solitons, in agreement with the formula $N_{soliton} = \lfloor (13\beta)^{-0.5} \rfloor$ ($\lfloor . \rfloor$ for integer part) [34].

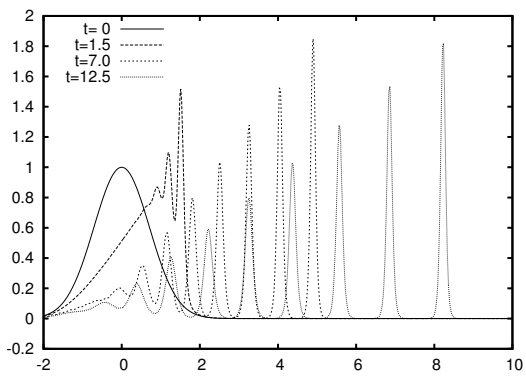


Fig. 6 Test-case 2: Solution at different times.

Computations have been carried out using $K = 240$ and $K = 300$ elements, with a polynomial degree $N = 5$. In terms of number of degrees of freedom, say dof, $K = 240$ and $K = 300$ correspond to the values $dof = 1200$ and $dof = 1500$ used

in [25] and in [57], respectively. The solution obtained at different times, without correcting step and for $K = 300$, is shown in Fig. 6. After the splitting into solitons one can check that the propagation velocity V of each soliton is proportional to its amplitude A, since the maxima are aligned on a straight line, and that as theoretically predicted $V = A/3$ [5].

Time	0.0	2.5	5.0	7.5	10.0	12.5
C_1	1.77245	1.77245	1.77245	1.77245	1.77245	1.77245
C_2	1.25331	1.25325	1.25231	1.25136	1.25174	1.25179
C_3	1.01957	1.01898	1.00940	1.00253	1.01040	1.01463

Table 3 Test-case 2: The invariant C_1, C_2 and C_3 at different times, $K = 240$ and $N = 5$.

Now we focus on the invariants and first provide results obtained without correcting step. Tables 3 and 4 give the invariants at several times, as obtained with $K = 240$ and $K = 300$ elements, respectively. As expected, the invariant C_1 is exactly conserved. Concerning the two other invariants, for $K = 240$, at the final time $t = 12.5$ the relative variations of the coefficients C_2 and C_3 are of $-0.12\,\%$ and $-0.48\,\%$, respectively. Of course, better results are obtained with $K = 300$ elements, with relative variations for C_2 and C_3 of $-0.04\,\%$ and $-0.13\,\%$.

Time	0.0	2.5	5.0	7.5	10.0	12.5
C_1	1.77245	1.77245	1.77245	1.77245	1.77245	1.77245
C_2	1.25331	1.25330	1.25335	1.25317	1.25283	1.25279
C_3	1.01957	1.01948	1.01912	1.01888	1.01864	1.01829

Table 4 Test-case 2: The invariant C_1, C_2 and C_3 at different times, $K = 300$ and $N = 5$.

The latter results can be compared to those in [57], obtained with a *quintic B-spline* FEM. With the same number of grid-points, the relative variations for C_1, C_2 and C_3 are $+0.05\,\%$, $+0.016\,\%$ and $+0.35\,\%$, respectively.

As explained in Section 3, the conservation of the two first invariants is obtained by using an additional correcting step, based on a projection on the manifold described by the two first invariants or by interpolation/extrapolation of two embedded RK IMEX schemes. The results obtained with $N = 5$ and $K = 240$ elements when using the two techniques are provided in Table 5. As expected, the overhead computational cost is negligible.

Time	0.0	2.5	5.0	7.5	10.0	12.5
C_1	1.77245	1.77245	1.77245	1.77245	1.77245	1.77245
C_2	1.25331	1.25331	1.25331	1.25331	1.25331	1.25331
C_3 (1)	1.01957	1.01907	1.01068	1.00415	1.01051	1.01660
C_3 (2)	1.01957	1.01910	1.01054	1.00250	1.00871	1.01662

Table 5 Test-case 2: The invariant C_1, C_2 and C_3 at different times, $K = 240$ and $N = 5$, SEM scheme completed with (1) a projection step and (2) an interpolation/extrapolation step.

4.3 Test-case 3:

Like in the pioneering paper [56] or *e.g.* more recently in [18], we consider the KdV equation with $\beta = 0.022^2$ in the periodic domain $(0,2)$ and assume the initial condition $u_0(x) = \cos(\pi x)$. The numerical solution is computed with $K = 160$ elements, a polynomial approximation degree $N = 5$ and a time step $\tau = 2.5\ 10^{-4}$. An additional correcting step based on interpolation/extrapolation is used, see Section 3.2.2.

Fig. 7 shows the numerical solution at different critical times $\{0,\ t_{\mathrm{B}} = 1/\pi \approx 0.3183,\ 3.6t_{\mathrm{B}} \approx 1.1459\}$, where $t_{\mathrm{B}} = -1/\min(u_0')$ is the so-called breakdown time, at which characteristics begin to cross for the Burgers equation. These results agree very well with the numerical results obtained with other schemes [18,58]. We observe that, at time t_{B}, the solution is about to breakdown and, at time $3.6t_{\mathrm{B}}$, we discern a train of 8 solitons.

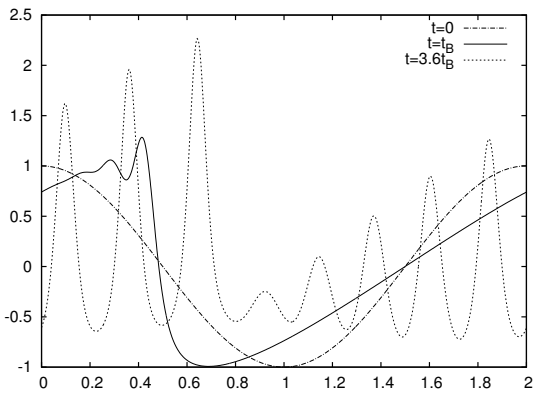


Fig. 7 Test-case 3: Numerical solution at times $t = \{0,\ t_{\mathrm{B}},\ 3.6t_{\mathrm{B}}\}$. Computation done with $K = 160$ elements, a polynomial degree $N = 5$ and a time step $\tau = 2.5\ 10^{-4}$.

As in [18], we then carry out the computation to the times $0.5t_{\mathrm{R}},\ t_{\mathrm{R}},\ 2t_{\mathrm{R}},\ 5t_{\mathrm{R}},\ 10t_{\mathrm{R}}$ and $20t_{\mathrm{R}}$, where $t_{\mathrm{R}} = 30.4t_{\mathrm{B}}$ is the so-called recurrence time, at which one expects to (approximately) recover the initial condition. The results obtained are presented in Fig. 8. These long time computations are stable, since no spurious oscillations appear. For the larger times, our results differ from the ones presented in [18]. One can check that such results are well converged in space and time: Fig. 9 shows a zoom on the interval $(1,1.6)$ of the numerical solutions at time $20t_{\mathrm{R}}$ obtained with different mesh sizes and time steps. The different results agree well apart from the one obtained on the coarser mesh. In this latter case, some little spurious oscillation appear (see Section 4.6). The contour levels of the numerical solution in the (x,t)-plane are plotted in Fig. 10 between times 0 and t_{R} at left, and between times $19t_{\mathrm{R}}$ and $20t_{\mathrm{R}}$ at right.

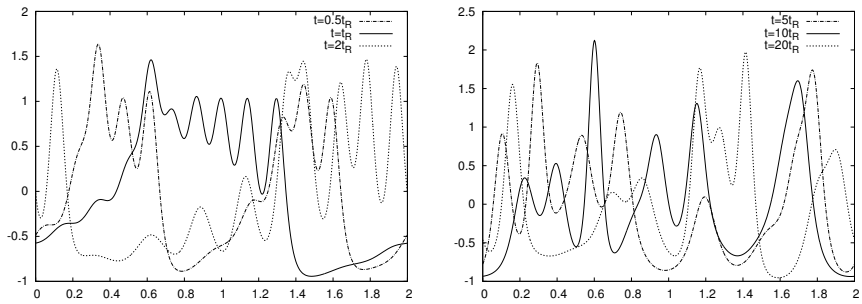


Fig. 8 Test-case 3: Numerical solution at times $t = 0.5t_{\mathrm{R}},\ t_{\mathrm{R}},\ 2t_{\mathrm{R}},\ 5t_{\mathrm{R}},\ 10t_{\mathrm{R}}$ and $20t_{\mathrm{R}}$. Computation done with $K = 160$ elements, a polynomial degree $N = 5$ and a time step $\tau = 2.5\ 10^{-4}$.

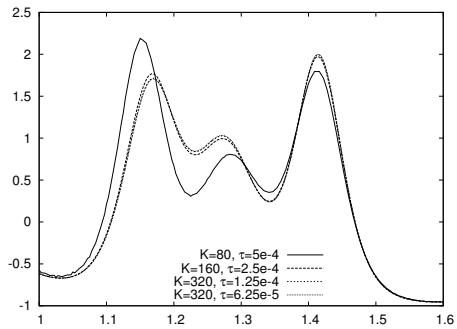


Fig. 9 Test-case 3: Zoom of the numerical solutions at times $20t_{\mathrm{R}}$. Computations done for various values of the number of elements K and time-step τ, and with a polynomial degree $N = 5$.

4.4 Test-case 4:

We consider the one-soliton solution of the KdV equation, with $\beta = 1$, on the whole line:

$$u(x,t) = \frac{12\kappa^2}{\cosh^2(\kappa x - 4\kappa^3 t)}. \tag{20}$$

This solution with $\kappa = 0.3$ is used to test the stability of our method in long time integration, as in [18]. In order to minimize the influence of boundary conditions, in our computation a large interval $(-80, 80)$ is chosen as computational domain, the exact solution being then close to the machine accuracy at the two end points. Then, periodic boundary conditions are used. We make comparison, after long time simulations, between the numerical solution and the "quasi-exact" solution obtained by periodization of (20). The simulation is performed with $K = 120$ cells, a polynomial degree $N = 5$ and a time step $\tau = 0.106667$ so that the number of degrees of freedom and the time step are comparable to those used in [18]. An additional correcting step based on interpolation/extrapolation is used, see Section 3.2.2. The results at time $t = 10^5/0.45$, $t = 2\ 10^5/0.45$ and $t = 10^6/0.45$ are displayed in Fig. 11. At these very long times (much larger than in [18]), one observes a small phase error whereas the shape is well preserved.

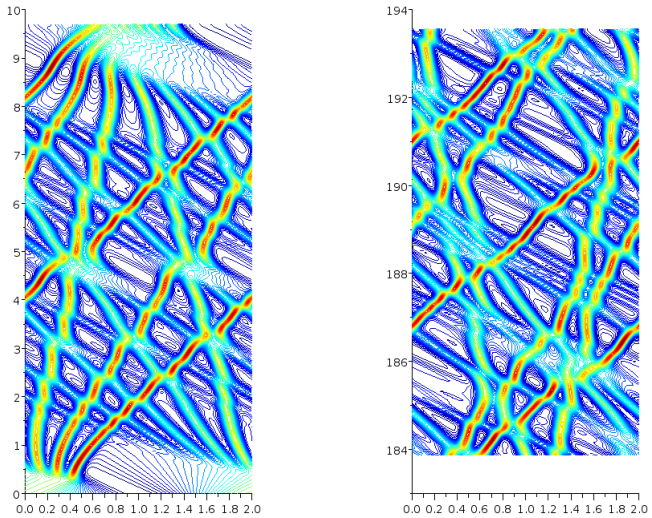


Fig. 10 Test-case 3: Contour levels of the numerical solution in the (x,t)-plane, between times 0 and t_{R} at left, between times $19t_{\mathrm{R}}$ and $20t_{\mathrm{R}}$ at right. Computation done with $K = 160$ elements, a polynomial degree $N = 5$ and a time step $\tau = 2.5\ 10^{-4}$.

4.5 Test-case 5

To address the non periodic case, as in [54] we compute the classical soliton solution

$$u_{ex}(x,t) = -2\mathrm{sech}^2(x-4t)$$

of the following KdV equation

$$\partial_t u - 3u\partial_x u + \partial_{xxx} u = 0.$$

The initial condition is given by the exact solution $u_{ex}(x,0)$ and we use the following boundary conditions

$$u(x_{\min},t) = b_-(t), \quad \partial_x u(x_{\max},t) = b'_+(t), \quad \text{and} \quad \partial_{xx} u(x_{\max},t) = b''_+(t),$$

where the boundary data are obtained from the exact solution. We use here the Strategy A presented in Section 2.2.1 to discretize the third order term but similar results have been obtained with the Strategy B. To compare our results with those obtained in [54], we first provide in Table 6 the numerical errors obtained at time $t = 0.5$ on the computational domain $(-10,12)$ for different mesh size, the time step being chosen small enough so that spatial errors dominate. The magnitude of the errors are comparable to those obtained in [54] and we observe optimal convergence order. Next, we carry out a simulation in the computational domain $(5,15)$ and up to time $t = 4.5$,

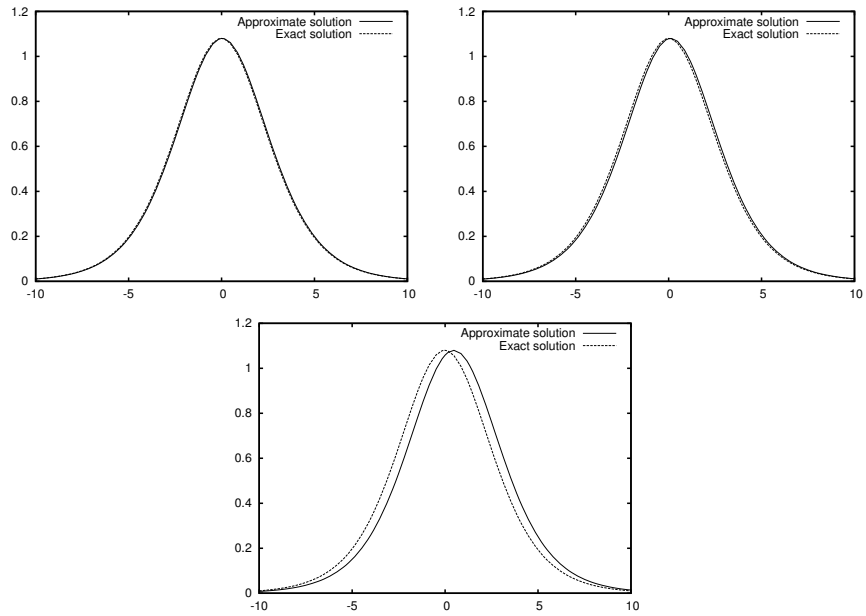


Fig. 11 Test-case 4: Zoom on numerical and quasi-exact solutions at times $t = 10^5/0.45$ (top left), $t = 2\,10^5/0.45$ (top right) and $t = 10^6/0.45$ (bottom). Computation done with $K = 120$ elements, a polynomial degree $N = 5$ and a time step $\tau = 0.106667$.

N		K=40	K=80		K=160		K=320	
		error	error	order	error	order	error	order
2	L^2	1.47e-2	1.06e-3	3.8	1.26e-04	3.1	1.55e-05	3.0
	L^∞	1.42e-2	1.41e-3	3.3	1.72e-04	3.0	2.15e-05	3.0
3	L^2	1.03e-3	4.79e-5	4.4	2.97e-6	4.0	1.85e-7	4.0
	L^∞	1.39e-3	9.49e-5	3.9	5.92e-6	4.0	3.72e-7	4.0

Table 6 Test-case 5: L^2 and L^∞ errors at time $t = 0.5$ with $(x_{\min}, x_{\max}) = (-10, 12)$.

using $N = 5$ and $K = 100$ elements. The results for different times are plotted in Figure 12 (only the values at the end points of the elements are visualized). At the initial time, the soliton is outside the computational domain at left. It goes through the whole domain from left to right during the simulation and is outside at right at the final time. We can observe that the boundary conditions are properly taken into account.

4.6 Underresolved test-case 2

The numerical results presented till now have been obtained using Strategy A to express the third order derivative and exact integration for the transport term, but close results can be obtained when using Strategy B, the SVV method instead of overinte-

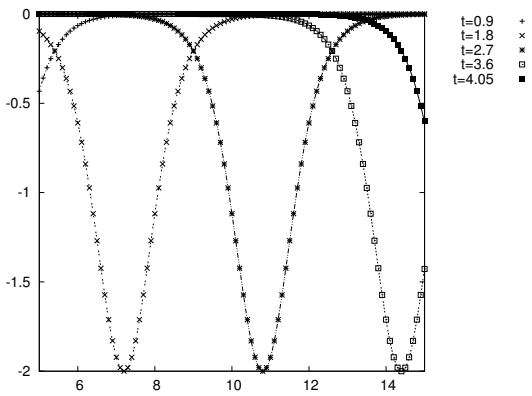


Fig. 12 Test-case 5: simulation in the computational domain $(5,15)$, numerical solution (points) and exact solution (lines) at different times. Computation done with $K = 100$ elements, a polynomial degree $N = 5$ and a time-step $\tau = 5\ 10^{-4}$.

gration or even with none of these treatments. To make comparisons between these different approaches, we consider here again the test-case 2 but with a rougher mesh, *i.e.* $K = 180$ elements, again with a polynomial degree $N = 5$. In time, one uses the IMEX scheme completed with an interpolation / extrapolation correcting step.

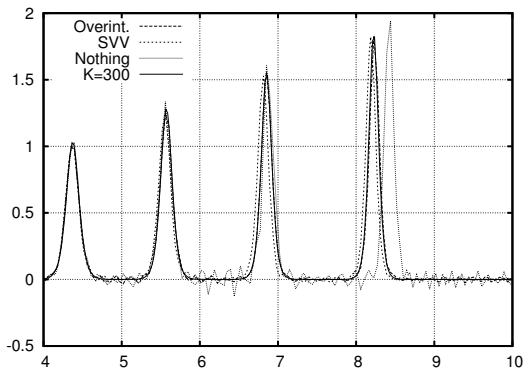


Fig. 13 Test-case 2: Underresolved solutions at time $t = 12.5$ with overintegration, SVV and none of them.

Fig. 13 compares, at the final time $t = 12.5$ and when using Strategy A, the results obtained (i) without overintegration, (ii) with exact integration of the non-linear term ($M = 8$) or (iii) if using the SVV method, with $m_N = [\sqrt{N}] = 2$ and $\varepsilon_N = h/N$, see Appendix A. For the sake of comparison, the solution obtained with $K = 300$ elements is also shown. Clearly, without overintegration or SVV stabilization, the solution shows oscillations and moreover the propagation velocities of the solitons are overvalued. The best result is here obtained with the overintegration of the non-linear term: The solution is indeed smoother and the propagation velocities are not

affected. One remarks that the SVV method slightly affects the propagation velocity. This phenomenon is strongly amplified when choosing $m_N = 0$, *i.e.* when also inserting viscosity on the low frequencies of the solution.

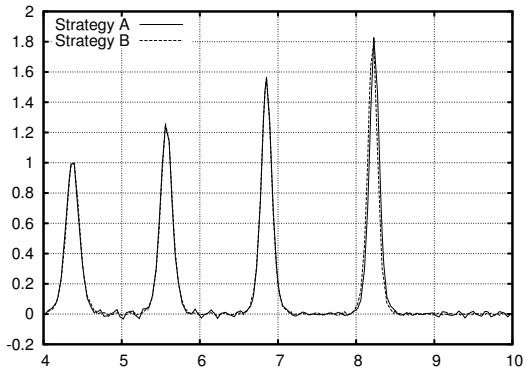


Fig. 14 Test-case 2: Underresolved solutions at time $t = 12.5$ with strategies A and B.

The strategies A and B are compared in Fig. 14, with exact integration of the transport term. One may observe that the propagation velocities are not affected, but that Strategy B yields qualitatively better results than Strategy A, since providing a smoother solution.

5 Concluding remarks

SEM schemes, *i.e.* involving high order C^0 continuous elements, together with IMEX RK time discretizations, and possibly an additional correction step for the preservation of invariants, have been proposed to address the approximation of dispersive equations and then applied to KdV. Such an approach is highly accurate and, differently to many other methods, offers a great flexibility in the choice of the approximation orders, both in space and in time. Moreover, the algorithms that have been described are simple, efficient and can be easily implemented in existing SEM softwares:

- The operators defined to address the high order derivative terms are easy to set up, since they are based on the usual mass matrix and usual stiffness (strategy A) or differentiation (strategy B) matrix. Taking advantage of the fact that the SEM matrix is diagonal, the present SEM schemes do not suffer from an increase of the number of unknowns.

- In time, we do not make use of a specific time scheme but instead propose choosing one of the family of IMEX RK schemes. This allows highly accurate implicit treatment of the linear term and explicit treatment of the convective term. The complexity of the time marching algorithm is then minimized and only submitted to the usual CFL constraint.

- Rather than using a specific time discretization, *e.g.* the Crank-Nicolson scheme as described in Section 3.2 (see Proposition 2), we have proposed two ways to enforce the exact conservation of the energy that are not specific to KdV or to 1D problems. Especially, the recently developed embedded IMEX RK schemes have appeared of great interest. They are easy to implement and cheap, since the secondary IMEX RK is just a minor modification of the main one.

- Also, it has been shown that an exact integration of the non-linear transport term was of interest.

Optimal accuracy results have been obtained using this methodology and difficult test problems have been successfully addressed : stiff problem, with long time integration, with effective boundary conditions.

These SEM schemes can be easily extended to multidimensional geometries and beyond the KdV equation, the present SEM methodology *a priori* applies to other PDEs showing higher order derivative terms. Thus, it would be of interest to address the Boussinesq system, *i.e.* the Saint-Venant system completed with additional dispersive terms, or more generally Boussinesq type equations, as *e.g.* done using a space-time LDG in [22]. Especially, one may think that the present SEM approach would be well adapted to handle a cross derivatives term, like $\partial_{txx}u$, by introducing the additional variable $f = \partial_{xx}u$ (strategy A). This will be investigated in near future.

A Spectral vanishing viscosity technique

The spectral vanishing viscosity (SVV) technique, as introduced in [38,46] for the Fourier and Legendre spectral approximations of the Burgers equation, is a viscous stabilization in spectral space. It is controlled by two parameters that depend on the polynomial approximation degree N, say m_N and ε_N, which define the threshold and the magnitude of an additional viscous term. In the frame of the SEM it is relevant to consider the stabilization term, in the reference interval $(-1,1)$

$$S_N = \varepsilon_N \partial_x Q_N(\partial_x u_N)$$

where we denote by u_N the polynomial approximation of degree N and by Q_N the spectral viscosity operator such that, for any polynomial $v_N \in \mathbb{P}_N(-1,1)$,

$$Q_N(v_N) = \sum_{k=0}^{N} \hat{Q}_k \hat{v}_k L_k(x)\,.$$

The $\hat{v}_k$ are the components of v_N in the (orthogonal and hierarchical) Legendre basis $\{L_k\}$, and the $\hat{Q}_k$ are monotonically increasing coefficients such that, $\hat{Q}_k = 0$ if $k \leqslant m_N$ and $0 < \hat{Q}_k \leqslant 1$ otherwise. As suggested in [38], we use $\hat{Q}_k = \exp(-(k-N)^2/(k-m_N)^2)$ for $m_N < k \leqslant N$.

Using now v_N as a test function, in weak form the stabilization term writes (boundary terms are neglected):

$$\begin{aligned}\int_{-1}^{1} v_N S_N\,dx &= \varepsilon_N \int_{-1}^{1} v_N \partial_x Q_N(\partial_x u_N)\,dx \\ &= -\varepsilon_N \int_{-1}^{1} \partial_x v_N Q_N(\partial_x u_N)\,dx \\ &= -\varepsilon_N \int_{-1}^{1} Q_N^{1/2}(\partial_x v_N) Q_N^{1/2}(\partial_x u_N)\,dx\end{aligned}$$

where in the last equality, which results from the orthogonality of the Legendre polynomials, $Q_N^{1/2}$ is defined as Q_N, but using the coefficients $\sqrt{\hat{Q}_k}$. This last form is especially of interest when multidimensional

SEM approximations are considered [51], whereas in the frame of the 1D SEM used in this paper, the two forms are strictly equivalent.

Acknowledgements This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement number 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission. We also thank our Colleagues D. Clamond and A. Galligo for fruitful discussions.

References

1. E.N. Aksan, A. Özdes, Numerical solution of the Korteweg-de Vries equation by Galerkin B-spline finite element method, *Appl. Math. Comput.* 175 (2006) 1256-1265.
2. U.M. Ascher, S.J. Ruth, B. Wetton, Implicit-explicit methods for time-dependent PDE's, *SIAM J. Numer. Anal.* 32 (1995) 797-823.
3. U.M. Ascher, S.J. Ruth, R.J. Spiteri, Implicit-explicit Runge-Kutta methods for time-dependent partial differential equations, *Appl. Num. Math.* 25 (1997) 151-167.
4. G.A. Baker, V.A. Dougalis, O.A. Karakashian, Convergence of Galerkin Approximations for the Korteweg-de Vries equation, *Mathematics of Computation* 40 (163) (1983) 419–433.
5. Y.U. Berezin, V.I. Karpman, Nonlinear evolution of disturbances in plasmas and other dispersive media, *Soviet Physics Jetp* 24 (5) (1967) 1049-1055.
6. J.L. Bona, V.A. Dougalis, O.A. Karakashian, W.R. McKinney, Conservative high-order numerical schemes for the generalized Korteweg-de Vries equation, *Phil. Trans. R. Soc. Lond.* 351 (1995) 107-164.
7. J.L. Bona, H. Chen, O. Karakashian, Y. Xing, Conservative, discontinuous Galerkin methods for the generalized Korteweg-de Vries equation, *Mathematics of Computation* 82 (283) (2013) 1401–1432.
8. N.D. Bueno, C. Mastroserio, Explicit methods based on a class of four stage fourth order Runge-Kutta methods for preserving quadratic invariants, *J. of Comput. and Appl. Math.* 20 (1996) 247-260.
9. J.C. Butcher, A histoty of Runge-Kutta methods, *Appl. Num. Math.* 25 (1997) 151-167.
10. M. Calvo, D. Hernández-Abreu, J.I. Montijano, L. Rández, On the preservation of invariants by explicit Runge-Kutta methods, *SIAM J. Sci. Comput.* 28(3) (2006), 868885.
11. M. Calvo, M. P. Laburta, J.I. Montijano, L. Rández, Runge-Kutta projection methods with low dispersion and dissipation errors, *Adv Comput Math* 41 (2015) 231251.
12. C. D. Cantwell, D. Moxey, A. Comerford, A. Bolis, G. Rocco, G. Mengaldo, D. De Grazia, S. Yakovlev, J-E. Lombard, D. Ekelschot, B. Jordi, H. Xu, Y. Mohamied, C. Eskilsson, B. Nelson, P. Vos, C. Biotto, R. M. Kirby, S. J. Sherwin, Nektar++: An open-source spectral/hp element framework, *Comput Phys Commun*, 192 (2015) 205-219.
13. C. Canuto, M. Hussaini, A. Quarteroni, T. Zang, *Spectral Methods, Evolution to Complex Geometries and Applications to Fluid Dynamics*, Springer Verlag, 2007.
14. G.F. Carey, Y. Shen, Approximation of the KdV equation by least square finite elements, *J. Comput. Methods in Appl. Mech. and Engrg.* 93 (1991) 1-11.
15. E. Celledoni, R.I. McLachlan, D.I. McLaren, B. Owen, G.R.W. Quispel, W.M. Wright, Energy-preserving Runge-Kutta methods, *ESAIM: M2AN* 43 (2009) 645-649.
16. G. Cohen, P. Monk, Gauss point mass lumping schemes for Maxwell's equations, *Numer. Methods Partial Differential Equations*, 14 (1998), 63-88.
17. G.J. Cooper, Stability of runge-kutta methods for trajectory problems, *IMA Journal of Numerical Analysis*, 7(1) (1987), 1-13.
18. Y. Cui, D.K. Mao, Numerical method satisfying the first two conservation laws for the Korteweg-de Vries equation, *J. Comput. Phys.*, 227(1) (2007), 376-399.
19. M.O. Deville, P.F. Fischer, E.H. Mund, *High-order methods for incompressible fluid flow*, Cambridge University Press, New York, 2002.
20. Y. Dubois-Pelerin, V. van Kemenade, M. Deville, An Object-Oriented Toolbox for Spectral Element Analysis, *J. Sci. Comput.* 14 (1) (1999) 1-29.
21. D. Dutykh, T. Katsaounis, D. Mitsotakis, Finite volume methods for unidirectional dispersive wave models, *Int. J. for Numerical Methods in Fluids*, 71 (2013) 717-736.

22. M. Dumbser, M. Facchini, A space-time discontinuous Galerkin method for Boussinesq-type equations, *Applied Mathematics and Computation*, 272 (2016) 336-346.
23. P. F. Fischer, J. W. Lottes, S. G. Kerkemeier, nek5000 Web page, http://nek5000.mcs.anl.gov, (2008).
24. D. Funaro, *Spectral elements for transport-dominated equations*, Springer-Verlag, Berlin, 1997.
25. L.R.T. Gardner, G.A. Gardner, A.H.A. Ali, Simulations of solitons using quadratic spline finite elements, *J. Comput. Methods in Appl. Mech. and Engrg.* 92 (1991) 231-243.
26. J.L. Guermond, R. Pasquetti, B. Popov, Entropy viscosity method for non-linear conservation laws, *J. Comput. Phys.* 230 (11) (2011) 4248-4267.
27. H. Holden, K.H. Karlsen, N.H. Risebro, Operator splitting methods for generalized Korteweg-de Vries equations, *J. Comput. Phys.* 153 (1999) 203-222.
28. T.J.R. Hugues, G.R. Feijoo, L. Mazzei, J.B. Quincy, The variational multiscale method - a paradigm for computational mechanics, *Comput. Methods in Appl. Mech. Engrg.* 166 (1998) 3-24.
29. A. Iserles, A. Zanna, Preserving algebraic invariants with Runge-Kutta methods, *J. of Comput. and Appl. Math.* 125 (2000) 69-81.
30. T. Kappeler, P. Topalov, Global wellposedness of KdV in $H^{-1}(\mathbb{T},\mathbb{R})$, *Duke Mathematical J.* 135 (2006) 327-360.
31. G.E. Karniadakis, S.J. Sherwin, *Spectral hp element methods for CFD*, Oxford Univ. Press, London, 1999.
32. R.M. Kirby, S.J. Sherwin, B. Cockburn, To CG or to HDG: A comparative study, *J. Sci. Comput.* 51 (2012) 183-212.
33. R.M. Kirby, G.E. Karniadakis, De-aliasing on non uniform grids: algorithms and applications, *J. Comput. Phys.* 191 (2003) 249-264.
34. V.I. Karpman, An asymptotic solution of the Korteweg-de Vries equation, *Physics Letters* 25A (10) (1967) 708-709.
35. P.D. Lax, C.D. Levermore, The small dispersion limit of the Korteweg-de Vries equation. I, *Comm. on Pure and Appl. Math.* XXXVI (1983) 253-290.
36. D. Levy, C.W. Shu, J. Yan, Local discontinuous Galerkin methods for nonlinear dispersive equations, *J. of Comput. Phys.* 196 (2004) 751-772.
37. Y. Maday, A.T. Patera, Spectral element methods for the incompressible Navier-Stokes equations, in A.K. Noor (Editor), State-of-the-Art Surveys in Computational Mechanics, ASME, New York, 1989, 71-143.
38. Y. Maday, S.M.O. Kaber, E. Tadmor, Legendre pseudo-spectral viscosity method for nonlinear conservation laws, *SIAM J. Numer. Anal.* 30 (2) (1993) 321-342.
39. J.W. Miles, The Korteweg-de Vries equation: a historical essay, *J. Fluid Mech.* 106 (1984) 131-147.
40. J.P. Ohlsson, P. Schlatter, P.F. Fischer, D.S. Henningson, Stabilization of the spectral element method in turbulent flow simulations, in *Lecture Notes in Computational Science and Engineering* 76, Springer-Verlag Berlin Heidelberg 2011, 449-458.
41. L. Pareschi, G. Russo, Implicit-Explicit Runge-Kutta schemes for stiff systems of differential equations, Recent trends in numerical analysis, 269-288, Nova Science Publishers, Inc. Commack, NY, USA 2000.
42. A. Samii, N. Panda, C. Michoski, C. Dawson, A hybridized discontinuous Galerkin method for the nonlinear Korteweg-de Vries equation, *J. Sci. Comput.* 68 (2016) 191-212.
43. J.M. Sanz-Serna, I. Christie, Petrov-Galerkin Methods for Nonlinear Dispersive Waves, *J. Comput. Phys.* 39 (1981) 94-102.
44. J.M. Sanz-Serna, An explicit finite-difference scheme with exact conservation properties, *J. Comput. Phys.* 47 (1982) 199-210.
45. J. Shen, A new dual-Petrov-Galerkin method for third and higher odd-order differential equations: Application to the KdV equation, *SIAM J. Numer. Anal.* 41 (5) (2003) 1595-1619.
46. E. Tadmor, Convergence of spectral methods for nonlinear conservation laws, *SIAM J. Numer. Anal.* 26 (1) (1989) 30-44.
47. T.R. Taha, M.J. Ablowitz, Analytical and numerical aspects of certain nonlinear evolution equations. III. Numerical, Korteweg-de Vries equation, *J. Comput. Phys.* 55 (1984) 231-253.
48. P.E.J. Vos, S.J. Sherwin, R.M. Kirby, From h to p efficiently: Implementing finite and spectral / hp element methods to achieve optimal performance for low- and high-order discretisations, *J. Comput. Phys.* 229 (2010) 5161-5181.
49. M. Walkley, M. Berzins, A finite element method for the one-dimensional extended Boussinesq equations, *Int. J. Numer. Meth. Fluids*, 29 (1999) 143-157.
50. R. Winther, A conservative finite element method for the Korteweg-de Vries equation, *Mathematics of computation*, 34 (149) (1980) 23-43.

51. C.J. Xu, R. Pasquetti, Stabilized spectral element computations of high Reynolds number incompressible flows, *J. Comput. Phys.* 196 (2) (2004) 680-704.
52. Y. Xu, C.W. Shu, Optimal error estimates of the semidiscrete local discontinuous Galerkin methods for high order wave equations, *SIAM J. Numer. Anal.* 50 (1) (2012) 79-104.
53. J. Yan, C.-W. Shu, Local discontinuous Galerkin methods for partial differential equations with higher order derivatives, *J. Sci. Comput.* 17 (1-4) (2002) 27-47.
54. J. Yan, C.-W. Shu, A Local Discontinuous Galerkin Method for KdV Type Equations, *SIAM J. Numer. Anal.* 40 (2) (2002) 769-791.
55. N. Yi, Y. Huang, H. Liu, A direct discontinuous Galerkin method for the generalized Korteweg-de Vries equation: Energy conservation and boundary effect. *J. Comput. Phys.* 242 (2013) 351-366.
56. N.J. Zabusky, M.D. Kruskal, Interaction of solitons in a collisionless plasma and the recurence of initial states, *Phys. Rev. Letters* 15 (6) (1965) 240-243.
57. S.I. Zaki, A quintic B-spline finite elements scheme for the KdVB equation, *J. Comput. Methods in Appl. Mech. and Engrg.* 188 (2000) 121-134.
58. P. Zhao, M. Qin, Multisymplectic geometry and multisymplectic Preissmann scheme for the KdV equation, J. Phys. A: Math. Gen. 33 (2000) 3613-3626.